

NONPARAMETRIC ESTIMATION OF THE JUMP RATE FOR NON-HOMOGENEOUS MARKED RENEWAL PROCESSES

ROMAIN AZAÏS, FRANÇOIS DUFOUR, AND ANNE GÉGOUT-PETIT

ABSTRACT. This paper is devoted to the nonparametric estimation of the jump rate and the cumulative rate for a general class of non-homogeneous marked renewal processes, defined on a separable metric space. In our framework, the estimation needs only one observation of the process within a long time. Our approach is based on a generalization of the multiplicative intensity model, introduced by Aalen in the seventies. We provide consistent estimators of these two functions, under some assumptions related to the ergodicity of an embedded chain and the characteristics of the process. The paper is illustrated by a numerical example.

1. INTRODUCTION

The purpose of this paper is to develop a nonparametric method for estimating the jump rate of a non-homogeneous marked renewal process, when only one observation of the process within a long time is available. Our estimation procedure is premised on a generalization of the well-known multiplicative intensity model, investigated by Aalen in [2, 3].

We introduce a general class of non-homogeneous marked renewal processes (NHMRP's), defined on an open subset of a separable metric space. The motion of the process depends on three characteristics namely the jump rate λ , which specifies the interarrival times, the transition kernel Q , and a function t^* , which plays the role of a deterministic censorship depending on the state of the process. In this framework, the jump rate λ is a function of two variables: a spatial mark and time. Here, our aim is to propose a nonparametric method for estimating both the jump rate λ and the cumulative rate from only one observation of the process within a long time interval. In addition, the class of non-homogeneous marked renewal processes which we consider may be related to particular piecewise-deterministic Markov processes (PDMP's, see the book [13]), whose transition kernel does not depend on time. This paper is also a keystone of [6], in which we provide a consistent nonparametric estimator of the conditional density associated to the jump rate of a general PDMP.

Aalen suggested in the middle of the seventies the famous multiplicative intensity model (see his PhD thesis [1], or [2, 3]). In this work [1, 2, 3], it is assumed that the intensity of the underlying counting process N can be written as the product of a predictable process Y

2010 *Mathematics Subject Classification.* Primary: 62G05, Secondary: 62M09.

Key words and phrases. Non-homogeneous marked renewal process, nonparametric estimation, jump rate estimation, Nelson-Aalen estimator, asymptotic consistency, ergodicity of Markov chains.

This work was supported by ARPEGE program of the French National Agency of Research (ANR), project "FAUTOCOES", number ANR-09-SEGI-004.

and a deterministic function λ , called jump rate or hazard rate. In this framework, the so-called Nelson-Aalen estimator provides an estimate of the cumulative rate $\Lambda(t) = \int_0^t \lambda(s)ds$. Ramlau-Hansen suggested a few years later, in [23], a nonparametric method for estimating directly the rate λ , by smoothing the Nelson-Aalen estimator by kernel methods.

A large number of estimation problems in survival analysis or in statistics of processes are related to counting processes depending on a spatial variable. This one may be seen as a mark or as a covariate. In this context, Aalen's estimator is proved to be a very popular and powerful method, since the multiplicative assumption is satisfied in a large variety of applications (see for instance [5]). In particular, one can apply the Nelson-Aalen approach for estimating the jump rate of a marked counting process, whose state space is finite, from a large number of independent observations. More recently in 2011, Comte *et al.* proposed in [10] an adaptive method for estimating the jump rate of a marker-dependent counting process, under the multiplicative assumption.

There exists an extensive literature on nonparametric and semiparametric estimation methods when the spatial mark belongs to a continuous state space. However, we do not attempt to present an exhaustive survey on this topic. A significant list of references on this research field can be found in [4, 5, 16, 20] and the references therein. In particular, McKeague and Utikal were interested in [21] in the estimation of the jump rate when the covariate takes its values in $[0, 1]$. Their approach is based on smoothing a Nelson-Aalen type estimator both in spatial and time directions. In particular, they prove the consistency of their estimator. Li and Doss chose another approach, based on a local linear fit in the spatial direction (see [19]). They extended McKeague and Utikal's work for the multidimensional case, and proved weak convergence results. One may also refer to the papers written by Utikal [25, 26] about jump rate estimation for two special classes of marked counting processes, under some continuous-time martingale assumptions. The Euclidean structure of the covariate state space is a keystone of the papers mentioned above. At the same time, nonparametric approaches have also been considered by Beran in [8], Stute in [24] and Dabrowska in [12], but for independent observations. Semiparametric methods have also been considered by many authors, beginning with Cox in [11]. The interested reader may consult the book [5] and the references therein for a complete review of the literature on these models.

In many aspects, our approach and the results mentioned above are different and complementary. These differences may be briefly described as follows. Our paper is based on a generalization of the multiplicative intensity model, involving a discretization of the state space, and, as a consequence, an approximation of the functions of interest. Indeed, we do not impose any conditions on the state space, such as to be Euclidean. This notably excludes the methods investigated by McKeague and Utikal, Li and Doss, or Utikal. Furthermore, these authors consider some assumptions about both continuous-time martingale properties and the asymptotic behavior of Y . From a practical point of view, this kind of assumption is not completely satisfactory due to the fact that these conditions may be difficult to check, especially in our case. On our side, we overcome this difficulty by providing tractable conditions, directly related to the primitive data of the process, to ensure the consistency of our estimator.

The present paper is divided into two parts. In the first one, we consider that the transition kernel only charges a finite set of points. It amounts to considering the state space is a discrete one. In this context, Theorem 3.1 states that the multiplicative model is

satisfied. In the second part, we assume that the kernel Q is diffuse, that is to say, it does not charge singletons. If the Z_i 's denote the marks of the underlying process, for any x and i , the indicator function $\mathbf{1}_{\{Z_i=x\}}$ is almost surely null. This rules out the method developed in the discrete case. Our procedure relies on a partition of the state space, labeled (A_k) . In this context, it appears intuitive to consider the counting process

$$N_n(A_k, t) = \sum_{i=0}^{n-1} \mathbf{1}_{\{Z_i \in A_k\}} \mathbf{1}_{\{S_{i+1} \leq t\}},$$

where the S_i 's denote the interarrival times. Although Aalen's multiplicative model does not hold for this counting process, the stochastic intensity of $N_n(A_k, t)$ is almost surely equivalent to the product $Y_n(A_k, t)l(A_k, t)$ (see Proposition 4.12), where $l(A_k, t)$ is an approximation of the jump rate $\lambda(x, t)$, for $x \in A_k$, and

$$Y_n(A_k, t) = \sum_{i=0}^{n-1} \mathbf{1}_{\{Z_i \in A_k\}} \mathbf{1}_{\{S_{i+1} \geq t\}}.$$

In this context, it is natural to introduce the following processes,

$$\widehat{L}_n(A_k, t) = \int_0^t Y_n(A_k, s)^+ dN_n(A_k, s),$$

where $Y_n(A_k, t)^+$ is the generalized inverse of $Y_n(A_k, t)$, and

$$L_n^*(A_k, t) = \int_0^t l(A_k, s) \mathbf{1}_{\{Y_n(A_k, s) > 0\}} ds,$$

In Aalen's papers, the difference $\widehat{L}_n(A_k, t) - L_n^*(A_k, t)$ is a continuous-time martingale, whereas on our part, it is not the case since there exists an extra-term $a_n(t)$, which vanishes when n goes to infinity. Intuitively, this means that the multiplicative model asymptotically makes sense. Referring to both Lengart's inequality and the asymptotic behavior of the extra-term $a_n(t)$, we prove that $\widehat{L}_n(A_k, t)$ is a consistent estimator of $L(A_k, t) = \int_0^t l(A_k, s) ds$ (see Proposition 4.16). We deduce from this a consistent estimator of the cumulative rate $\Lambda(x, t) = \int_0^t \lambda(x, s) ds$ (see Theorem 4.21), since $L(A_k, t)$ and $\Lambda(x, t)$ are close for $x \in A_k$ (see Lemma 4.18). Next, we focus on smoothing this estimator by kernel methods in order to suggest a consistent estimator of $l(A_k, t)$ (see Proposition 4.23) and, therefore, of $\lambda(x, t)$ (see Theorem 4.24). An inherent difficulty throughout this paper is related to the presence of the deterministic censorship.

The paper is organized in the following way. We first give, in Section 2, the precise definition of the class of non-homogeneous marked renewal processes which we are interested in, and we provide an example of application in reliability. We state also some technical results about continuous-time martingales and conditional independences. Section 3 is devoted to the discrete case, where we consider that the transition kernel Q only charges a finite number of points. The main contribution of the paper lies in Section 4, in which we do not impose any conditions on the state space. In this part, we provide consistent estimators of the cumulative rate (see Theorem 4.21) and the jump rate (see Theorem 4.24). Finally in Section 5, we present a numerical example to illustrate the good behavior of our estimators on finite sample size.

2. DEFINITION AND FIRST RESULTS

Let E be an open subset of a separable metric space (\mathcal{E}, d) . One considers a Markov chain $(Z_n)_{n \geq 0}$ on $(E, \mathcal{B}(E))$, defined on a probability space $(\Omega, \mathcal{A}, \mathbf{P}_{\nu_0})$. The distribution of Z_0 is assumed to be ν_0 . An equivalent formulation (see for instance Theorem 2.4.3 of [7]) is given by,

$$(1) \quad \forall n \geq 1, Z_n = \psi(Z_{n-1}, \varepsilon_{n-1}),$$

where ψ is a measurable function, and $(\varepsilon_n)_{n \geq 0}$ is a sequence of independent and identically distributed random variables. Let $\lambda : E \times \mathbf{R}_+ \rightarrow \mathbf{R}_+$ be a measurable function and Φ a deterministic flow on E . One may associate to Φ the deterministic exit time from E :

$$(2) \quad \forall \xi \in E, t^*(\xi) = \inf\{t > 0 : \Phi(\xi, t) \in \partial E\}.$$

One remarks that the function t^* may take the value $+\infty$. Now, one defines a sequence $(S_n)_{n \geq 1}$ on $(\mathbf{R}_+, \mathcal{B}(\mathbf{R}_+))$. The distribution of S_1 is given from its survival function,

$$\begin{aligned} \forall t \geq 0, \mathbf{P}_{\nu_0}(S_1 > t | \{Z_i : i \geq 0\}) &= \mathbf{P}_{\nu_0}(S_1 > t | Z_0) \\ &= \exp\left(-\int_0^t \lambda(Z_0, s) ds\right) \mathbf{1}_{\{0 \leq t < t^*(Z_0)\}}. \end{aligned}$$

For each integer $n \geq 2$, the distribution of S_n satisfies, for any $t \geq 0$,

$$\begin{aligned} (3) \quad \mathbf{P}_{\nu_0}(S_n > t | \{Z_i : i \geq 0\}, S_1, \dots, S_{n-1}) &= \mathbf{P}_{\nu_0}(S_n > t | Z_{n-1}) \\ &= \exp\left(-\int_0^t \lambda(Z_{n-1}, s) ds\right) \mathbf{1}_{\{0 \leq t < t^*(Z_{n-1})\}}. \end{aligned}$$

Hence, there exist a function φ and a sequence of independent and identically distributed random variables $(\delta_n)_{n \geq 0}$, which is independent of the sequence $(\varepsilon_n)_{n \geq 0}$, such that,

$$\forall n \geq 1, S_n = \varphi(Z_{n-1}, \delta_{n-1}).$$

One assumes that the two sequences $(\varepsilon_n)_{n \geq 0}$ and $(\delta_n)_{n \geq 0}$ are independent of Z_0 . The underlying non-homogeneous renewal process $(X_t)_{t \geq 0}$ is defined by,

$$\forall t \geq 0, X_t = Z_n \quad \text{if} \quad S_0 + \dots + S_n \leq t < S_0 + \dots + S_{n+1},$$

with the usual convention $S_0 = 0$.

We recall that we are interested here in the nonparametric estimation of the jump rate λ , from one observation of the embedded chain $(Z_n, S_n)_{n \geq 0}$. This class of non-homogeneous renewal models may be related to PDMP's, for which the transition kernel Q does not depend on time. Hence, the estimation method developed in this paper is very useful in order to estimate the conditional distribution of the interarrival times for PDMP's (see [6]). Nevertheless, providing a method for estimating the jump rate for this class of stochastic models has an intrinsic interest. In the following, we present an example in reliability of non-homogeneous marked renewal process satisfying the model mentioned above.

Let us consider a machine, whose production configuration takes its values in an open subset of \mathbf{R}^d . The dynamic of the regime is assumed to be a non-homogeneous renewal process: the state is piecewise-constant until a failure spontaneously occurs. One naturally considers that the failure rate depends on the production regime. When the machine breaks down, the repair occurs instantaneously, and the machine configuration is randomly changed, according to a transition kernel Q depending only on the previous working state.

In addition, one may consider that there exists a deterministic period of inspection, depending on the production configuration too. The inspection is instantaneous and the next regime changes according to the kernel Q . The estimation of the failure rate from only one observation of the regime state within a long time may bring some informations about the behavior of the production machine. The main benefit of this approach is as follows: the estimation does not need the observation of a great number of similar machines.

One may associate to the jump rate λ , the cumulative rate, the survival function and the probability density function, which are related to it. The conditional density f satisfies,

$$(4) \quad \forall \xi \in E, \quad \forall t \geq 0, \quad f(\xi, t) = \lambda(\xi, t) \exp \left(- \int_0^t \lambda(\xi, s) ds \right).$$

The cumulative rate Λ is defined by,

$$(5) \quad \forall \xi \in E, \quad \forall t \geq 0, \quad \Lambda(\xi, t) = \int_0^t \lambda(\xi, s) ds.$$

Finally, the conditional survival function is denoted by G .

$$(6) \quad \forall \xi \in E, \quad \forall t \geq 0, \quad G(\xi, t) = \exp \left(- \int_0^t \lambda(\xi, s) ds \right).$$

For each $n \geq 1$, \mathcal{G}_n denotes the σ -field generated by the n first Z_i 's,

$$(7) \quad \mathcal{G}_n = \sigma(Z_0, \dots, Z_{n-1}).$$

For each integer i , the one-jump counting process N^{i+1} is given for any $t \geq 0$, by

$$N^{i+1}(t) = \mathbf{1}_{\{S_{i+1} \leq t\}},$$

and $(\mathcal{F}_t^{i+1})_{t \geq 0}$ denotes the associated filtration. In this section, we shall prove two results: the first one is related to two conditional independence properties; the second one deals with the continuous-time martingale associated to the counting process N^{i+1} in a special filtration.

Proposition 2.1. *Let n be an integer and $1 \leq i \leq n$. For each integer $j \neq i$, let $t_j \geq 0$ and $t \geq 0$. Then, we have*

$$\bigvee_{j \neq i} \mathcal{F}_{t_j}^j \perp_{\mathcal{G}_n} \mathcal{F}_t^i \quad \text{and} \quad \mathcal{F}_t^i \perp_{\sigma(Z_{i-1})} \mathcal{G}_n.$$

Furthermore, we deduce from Proposition 6.8 of [17] this immediate corollary: for any $s < t$, $0 \leq i \leq n-1$,

$$\bigvee_{j \neq i+1} \mathcal{F}_s^j \perp_{\mathcal{G}_n \vee \mathcal{F}_s^{i+1}} \mathcal{F}_t^{i+1} \quad \text{and} \quad \mathcal{F}_t^{i+1} \perp_{\sigma(Z_i) \vee \mathcal{F}_s^{i+1}} \mathcal{G}_n.$$

Proof. We state the first conditional independence. Let us consider h_1, \dots, h_n some bounded measurable functions mapping from \mathbf{R}_+ to \mathbf{R} . From (3), we deduce

$$\begin{aligned} & \mathbf{E}_{\nu_0} [h_1(S_1) \dots h_n(S_n) | \mathcal{G}_n] \\ &= \mathbf{E}_{\nu_0} [h_1(S_1) \dots h_{n-1}(S_{n-1}) \mathbf{E}_{\nu_0} [h_n(S_n) | \mathcal{G}_n \vee \sigma(S_1, \dots, S_{n-1})] | \mathcal{G}_n] \\ &= \mathbf{E}_{\nu_0} [h_1(S_1) \dots h_{n-1}(S_{n-1}) \mathbf{E}_{\nu_0} [h_n(S_n) | \mathcal{G}_n] | \mathcal{G}_n] \\ (8) \quad &= \mathbf{E}_{\nu_0} [h_1(S_1) \dots h_{n-1}(S_{n-1}) | \mathcal{G}_n] \mathbf{E}_{\nu_0} [h_n(S_n) | \mathcal{G}_n]. \end{aligned}$$

Moreover, by (1),

$$\mathbf{E}_{\nu_0}[h_1(S_1) \dots h_{n-1}(S_{n-1}) | \mathcal{G}_n] = \mathbf{E}_{\nu_0}[\mathbf{E}_{\nu_0}[h_1(S_1) \dots h_{n-1}(S_{n-1}) | \mathcal{G}_{n-1} \vee \sigma(\varepsilon_{n-2})] | \mathcal{G}_n].$$

The product $h_1(S_1) \dots h_{n-1}(S_{n-1})$ is $\mathcal{G}_{n-1} \vee \sigma(\delta_0, \dots, \delta_{n-2})$ -measurable and

$$\sigma(\varepsilon_{n-2}) \perp \mathcal{G}_{n-1} \vee \sigma(\delta_0, \dots, \delta_{n-2}).$$

Together with (3) [9, page 308], we obtain

$$\mathbf{E}_{\nu_0}[h_1(S_1) \dots h_n(S_{n-1}) | \mathcal{G}_{n-1} \vee \sigma(\varepsilon_{n-2})] = \mathbf{E}_{\nu_0}[h_1(S_1) \dots h_{n-1}(S_{n-1}) | \mathcal{G}_{n-1}].$$

Finally, we have

$$\mathbf{E}_{\nu_0}[h_1(S_1) \dots h_{n-1}(S_{n-1}) | \mathcal{G}_n] = \mathbf{E}_{\nu_0}[h_1(S_1) \dots h_{n-1}(S_{n-1}) | \mathcal{G}_{n-1}].$$

In the light of (8), we obtain

$$\mathbf{E}_{\nu_0}[h_1(S_1) \dots h_n(S_n) | \mathcal{G}_n] = \mathbf{E}_{\nu_0}[h_1(S_1) \dots h_{n-1}(S_{n-1}) | \mathcal{G}_{n-1}] \mathbf{E}_{\nu_0}[h_n(S_n) | \mathcal{G}_n].$$

Thus, a straightforward induction leads to

$$\mathbf{E}_{\nu_0}[h_1(S_1) \dots h_n(S_n) | \mathcal{G}_n] = \prod_{i=1}^n \mathbf{E}_{\nu_0}[h_i(S_i) | \mathcal{G}_i].$$

Furthermore, taking for $j \neq i$, $h_j = \mathbf{1}$, yields to

$$\mathbf{E}_{\nu_0}[h_i(S_i) | \mathcal{G}_n] = \mathbf{E}_{\nu_0}[h_i(S_i) | \mathcal{G}_i].$$

Hence,

$$\mathbf{E}_{\nu_0}[h_1(S_1) \dots h_n(S_n) | \mathcal{G}_n] = \prod_{i=1}^n \mathbf{E}_{\nu_0}[h_i(S_i) | \mathcal{G}_i].$$

Thus, we have

$$\bigvee_{j \neq i} \sigma(S_j) \perp_{\mathcal{G}_n} \sigma(S_i),$$

that immediately implies the expected result. The second conditional independence is straightforward from (3). \square

We have also a continuous-time martingale property. This is the one associated to the counting process N^{i+1} .

Lemma 2.2. *For each integer i , the process M^{i+1} given by,*

$$(9) \quad \forall 0 \leq t < t^*(Z_i), \quad M^{i+1}(t) = N^{i+1}(t) - \int_0^t \lambda(Z_i, u) \mathbf{1}_{\{S_{i+1} \geq u\}} du,$$

is a continuous-time martingale in the filtration $(\sigma(Z_i) \vee \mathcal{F}_s^{i+1})_{0 \leq s < t^(Z_i)}$.*

Proof. Let $0 \leq s < t < t^*(Z_i)$. In order to prove that M^{i+1} is a martingale, we have to show that

$$\begin{aligned}
 & \mathbf{E}_{\nu_0} \left[N^{i+1}(t) - \int_0^t \lambda(Z_i, u) \mathbf{1}_{\{S_{i+1} \geq u\}} du \mid \sigma(Z_i) \vee \mathcal{F}_s^{i+1} \right] = M^{i+1}(s) \\
 \Leftrightarrow & \mathbf{E}_{\nu_0} \left[\mathbf{1}_{\{s < S_{i+1} \leq t\}} - \int_s^t \lambda(Z_i, u) \mathbf{1}_{\{S_{i+1} \geq u\}} du \mid \sigma(Z_i) \vee \mathcal{F}_s^{i+1} \right] = 0 \\
 \Leftrightarrow & \mathbf{1}_{\{S_{i+1} > s\}} \mathbf{E}_{\nu_0} [\mathbf{1}_{\{S_{i+1} \leq t\}} \mid \sigma(Z_i) \vee \mathcal{F}_s^{i+1}] \\
 (10) \quad & = \mathbf{1}_{\{S_{i+1} > s\}} \mathbf{E}_{\nu_0} \left[\int_s^t \lambda(Z_i, u) \mathbf{1}_{\{S_{i+1} \geq u\}} du \mid \sigma(Z_i) \vee \mathcal{F}_s^{i+1} \right].
 \end{aligned}$$

First, we shall prove that (10) is equivalent to

$$\begin{aligned}
 & \mathbf{1}_{\{S_{i+1} > s\}} \mathbf{E}_{\nu_0} [\mathbf{1}_{\{S_{i+1} \leq t\}} \mid \sigma(Z_i) \vee \sigma(N^{i+1}(s))] \\
 (11) \quad & = \mathbf{1}_{\{S_{i+1} > s\}} \mathbf{E}_{\nu_0} \left[\int_s^t \lambda(Z_i, u) \mathbf{1}_{\{S_{i+1} \geq u\}} du \mid \sigma(Z_i) \vee \sigma(N^{i+1}(s)) \right].
 \end{aligned}$$

In the light of Lemma 6.2 of [17], if the following conditions are satisfied,

$$\{S_{i+1} > s\} \in \left(\sigma(Z_i) \vee \mathcal{F}_s^{i+1} \right) \cap \left(\sigma(Z_i) \vee \sigma(N^{i+1}(s)) \right)$$

and

$$\{S_{i+1} > s\} \cap \left(\sigma(Z_i) \vee \mathcal{F}_s^{i+1} \right) = \{S_{i+1} > s\} \cap \left(\sigma(Z_i) \vee \sigma(N^{i+1}(s)) \right),$$

then (10) and (11) are equivalent. The first condition obviously holds. On the other hand, as $\sigma(N^{i+1}(s))$ is a sub- σ -field of \mathcal{F}_s^{i+1} , we have

$$\{S_{i+1} > s\} \cap \left(\sigma(Z_i) \vee \mathcal{F}_s^{i+1} \right) \supset \{S_{i+1} > s\} \cap \left(\sigma(Z_i) \vee \sigma(N^{i+1}(s)) \right).$$

Hence, we only have to prove the reciprocal inclusion,

$$\{S_{i+1} > s\} \cap \left(\sigma(Z_i) \vee \mathcal{F}_s^{i+1} \right) \subset \{S_{i+1} > s\} \cap \left(\sigma(Z_i) \vee \sigma(N^{i+1}(s)) \right).$$

Let us consider the λ -system (see for instance Definition 1.10 of [18]) \mathcal{D} given by

$$\mathcal{D} = \left\{ C \in \sigma(Z_i) \vee \mathcal{F}_s^{i+1} : \{S_{i+1} > s\} \cap C \in \{S_{i+1} > s\} \cap \left(\sigma(Z_i) \vee \sigma(N^{i+1}(s)) \right) \right\},$$

and the π -system (see Definition 1.1 of [18]) \mathcal{C} defined by

$$\mathcal{C} = \left\{ \Omega, \{S_{i+1} \leq u\} \cap \{Z_i \in B\}, \{Z_i \in B\} : 0 \leq u \leq s, B \in \mathcal{B}(E) \right\}.$$

We immediately obtain that $\mathcal{C} \subset \mathcal{D}$ and $\sigma(\mathcal{C}) = \sigma(Z_i) \vee \mathcal{F}_s^{i+1}$. Moreover, $\sigma(Z_i) \vee \mathcal{F}_s^{i+1} \subset \mathcal{D}$ on the strength of the monotone class theorem (see for example Theorem 1.19 of [18]). From the definition of \mathcal{D} , we straightforward deduce the reciprocal inclusion. Consequently, (10) and (11) are equivalent. Now, we only have to verify that (11) holds. On the one hand, we have

$$\begin{aligned}
 & \mathbf{1}_{\{S_{i+1} > s\}} \mathbf{E}_{\nu_0} [\mathbf{1}_{\{S_{i+1} \leq t\}} \mid \sigma(Z_i) \vee \sigma(N^{i+1}(s))] \\
 & = \mathbf{1}_{\{S_{i+1} > s\}} \mathbf{P}_{\nu_0}(S_{i+1} \leq t \mid Z_i, S_{i+1} > s) \\
 & = \mathbf{1}_{\{S_{i+1} > s\}} \frac{\mathbf{P}_{\nu_0}(s < S_{i+1} \leq t \mid Z_i)}{\mathbf{P}_{\nu_0}(S_{i+1} > s \mid Z_i)}.
 \end{aligned}$$

From the definitions of f (4) and G (6), we deduce that

$$\begin{aligned} & \mathbf{1}_{\{S_{i+1} > s\}} \mathbf{E}_{\nu_0} [\mathbf{1}_{\{S_{i+1} \leq t\}} \mid \sigma(Z_i) \vee \sigma(N^{i+1}(s))] \\ &= \mathbf{1}_{\{S_{i+1} > s\}} \frac{\int_s^t f(Z_i, u) du}{G(Z_i, s)}. \end{aligned}$$

On the other hand,

$$\begin{aligned} & \mathbf{1}_{\{S_{i+1} > s\}} \mathbf{E}_{\nu_0} \left[\int_s^t \lambda(Z_i, u) \mathbf{1}_{\{S_{i+1} \geq u\}} du \mid \sigma(Z_i) \vee \sigma(N^{i+1}(s)) \right] \\ &= \mathbf{1}_{\{S_{i+1} > s\}} \int_s^t \lambda(Z_i, u) \mathbf{P}_{\nu_0}(S_{i+1} \geq u \mid Z_i, S_{i+1} > s) du \\ &= \mathbf{1}_{\{S_{i+1} > s\}} \int_s^t \lambda(Z_i, u) \frac{\mathbf{P}_{\nu_0}(S_{i+1} \geq u \mid Z_i)}{\mathbf{P}_{\nu_0}(S_{i+1} > s \mid Z_i)} du. \end{aligned}$$

Thus, since $u \leq t < t^*(Z_i)$,

$$\begin{aligned} & \mathbf{1}_{\{S_{i+1} > s\}} \mathbf{E}_{\nu_0} \left[\int_s^t \lambda(Z_i, u) \mathbf{1}_{\{S_{i+1} \geq u\}} du \mid \sigma(Z_i) \vee \sigma(N^{i+1}(s)) \right] \\ &= \mathbf{1}_{\{S_{i+1} > s\}} \frac{\int_s^t \lambda(Z_i, u) G(Z_i, u) du}{G(Z_i, s)} \\ &= \mathbf{1}_{\{S_{i+1} > s\}} \frac{\int_s^t f(Z_i, u) du}{G(Z_i, s)}. \end{aligned}$$

As a consequence, we proved (11) and the expected result. \square

Proposition 2.1 and Lemma 2.2 are prominent for the next results. In addition, the reference to the underlying probability measure \mathbf{P}_{ν_0} will be implicit in the text. For the sake of readability, we shall write \mathbf{P}_x instead of $\mathbf{P}_{\delta_{\{x\}}}$.

3. DISCRETE STATE SPACE

We assume here that the transition kernel Q only charges a finite set which we denotes $\{x_1, \dots, x_M\}$. One may associate to each x_i the deterministic exit time $t_i^* = t^*(x_i)$. Now, one considers that k is fixed. In this section, we shall prove in Theorem 3.1 that the multiplicative model is satisfied for estimating the cumulative rate $\Lambda(x_k, \cdot)$.

For each integer n , let us introduce the counting process $N_n(x_k, \cdot)$ by,

$$(12) \quad \forall t \geq 0, \quad N_n(x_k, t) = \sum_{i=0}^{n-1} \mathbf{1}_{\{S_{i+1} \leq t\}} \mathbf{1}_{\{Z_i = x_k\}}.$$

In addition, for any $t \geq 0$, we define

$$(13) \quad Y_n(x_k, t) = \sum_{i=0}^{n-1} \mathbf{1}_{\{S_{i+1} \geq t\}} \mathbf{1}_{\{Z_i = x_k\}}.$$

Theorem 3.1. *Let $n \geq 1$. The process $M_n(x_k, \cdot)$ defined by,*

$$(14) \quad \forall 0 \leq t < t_k^*, \quad M_n(x_k, t) = N_n(x_k, t) - \int_0^t \lambda(x_k, s) Y_n(x_k, s) ds$$

is a $(\tilde{\mathcal{F}}_t^n)_{0 \leq t < t_k^*}$ -continuous-time martingale under \mathbf{P}_{ν_0} , with,

$$\forall 0 \leq t < t_k^*, \quad \tilde{\mathcal{F}}_t^n = \mathcal{G}_n \vee \bigvee_{i=0}^{n-1} \mathcal{F}_t^{i+1}.$$

Proof. Let $0 \leq s < t < t_k^*$. Plugging (12) and (13) in (14), we have

$$\begin{aligned} M_n(x_k, t) &= \sum_{i=0}^{n-1} \left\{ N^{i+1}(t) \mathbf{1}_{\{Z_i=x_k\}} - \int_0^t \lambda(Z_i, u) \mathbf{1}_{\{S_{i+1} \geq u\}} \mathbf{1}_{\{Z_i=x_k\}} du \right\} \\ (15) \quad &= \sum_{i=0}^{n-1} M^{i+1}(s) \mathbf{1}_{\{Z_i=x_k\}}, \end{aligned}$$

by (9). Thus, we obtain

$$\mathbf{E}_{\nu_0}[M_n(x_k, t) | \tilde{\mathcal{F}}_s^n] = \sum_{i=0}^{n-1} \left\{ \mathbf{E}_{\nu_0}[N^{i+1}(t) | \tilde{\mathcal{F}}_s^n] - \mathbf{E}_{\nu_0} \left[\int_0^t \mathbf{1}_{\{S_{i+1} \geq u\}} \lambda(x_k, u) du | \tilde{\mathcal{F}}_s^n \right] \right\} \mathbf{1}_{\{Z_i=x_k\}}.$$

On the strength of Proposition 2.1, we have

$$\bigvee_{j \neq i+1} \mathcal{F}_s^j \perp_{\mathcal{G}_n \vee \mathcal{F}_s^{i+1}} \mathcal{F}_t^{i+1}.$$

Together with Proposition 6.6 of [17], we deduce

$$\begin{aligned} &\mathbf{E}_{\nu_0}[M_n(x_k, t) | \tilde{\mathcal{F}}_s^n] \\ &= \sum_{i=0}^{n-1} \left\{ \mathbf{E}_{\nu_0}[N^{i+1}(t) | \mathcal{G}_n \vee \mathcal{F}_s^{i+1}] - \mathbf{E}_{\nu_0} \left[\int_0^t \mathbf{1}_{\{S_{i+1} \geq u\}} \lambda(x_k, u) du | \mathcal{G}_n \vee \mathcal{F}_s^{i+1} \right] \right\} \mathbf{1}_{\{Z_i=x_k\}}. \end{aligned}$$

Furthermore, from Proposition 2.1, we have

$$\mathcal{F}_t^{i+1} \perp_{\sigma(Z_i) \vee \mathcal{F}_s^{i+1}} \mathcal{G}_n.$$

Thus, in the light of Proposition 6.6 of [17] again, we have

$$\begin{aligned} \mathbf{E}_{\nu_0}[M_n(x_k, t) | \tilde{\mathcal{F}}_s^n] &= \sum_{i=0}^{n-1} \left\{ \mathbf{E}_{\nu_0}[N^{i+1}(t) | \sigma(Z_i) \vee \mathcal{F}_s^{i+1}] \mathbf{1}_{\{Z_i=x_k\}} \right. \\ &\quad \left. - \mathbf{E}_{\nu_0} \left[\int_0^t \mathbf{1}_{\{S_{i+1} \geq u\}} \lambda(x_k, u) du | \sigma(Z_i) \vee \mathcal{F}_s^{i+1} \right] \mathbf{1}_{\{Z_i=x_k\}} \right\}. \end{aligned}$$

By Lemma 2.2, this yields to

$$\mathbf{E}_{\nu_0}[M_n(x_k, t) | \tilde{\mathcal{F}}_s^n] = \sum_{i=0}^{n-1} M^{i+1}(s) \mathbf{1}_{\{Z_i=x_k\}}.$$

Finally, together with (15), $M_n(x_k, \cdot)$ is, therefore, a martingale. \square

Theorem 3.1 states that one may estimate the cumulative rate $\Lambda(x_k, t)$ with the Nelson-Aalen estimator $\hat{\Lambda}_n(x_k, t)$ given by

$$\hat{\Lambda}_n(x_k, t) = \sum_{i=0}^{n-1} \frac{1}{Y_n(x_k, S_{i+1})} \mathbf{1}_{\{S_{i+1} \leq t\}} \mathbf{1}_{\{Z_i=x_k\}}.$$

One refers the interested reader to [2, 3] or [5] for a general survey on the properties of this estimator.

4. CONTINUOUS STATE SPACE

The present section is divided into two parts. In the first one, we provide an estimate of the cumulative rate Λ . The second part deals with the estimation of the jump rate λ by smoothing the estimator of Λ by kernel methods.

4.1. Estimation of Λ . Let us assume that the transition kernel Q is diffuse, that is to say, Q does not charge singletons. The previous procedure is ruled out, since for any $x \in E$ and each integer i , $\mathbf{1}_{\{Z_i=x\}} = 0$ almost surely. As a consequence, we shall naturally approximate under regularity conditions the jump rate in x , by the jump rate given the state is in a neighborhood of x . As mentioned in the introduction, we shall consider the counting process $N_n(A, t)$ defined for $A \in \mathcal{B}(E)$ by

$$N_n(A, t) = \sum_{i=0}^{n-1} \mathbf{1}_{\{Z_i \in A\}} \mathbf{1}_{\{S_{i+1} \leq t\}}.$$

We will see (27) that the stochastic intensity of $N_n(A, t)$ in a well-chosen filtration is

$$\sum_{i=0}^{n-1} \mathbf{1}_{\{Z_i \in A\}} \mathbf{1}_{\{S_{i+1} \geq t\}} \lambda(Z_i, t).$$

The first results of this section deal with the ergodicity of the underlying Markov chains. Their properties are prominent to establish the asymptotic behavior of the stochastic intensity of $N_n(A, t)$ (see Proposition 4.12). Based on these results, we will study in Proposition 4.16 and Theorem 4.17 the estimation of $L(A, t) = \int_0^t l(A, s) ds$, where $l(A, t)$ is an approximation of the jump rate $\lambda(x, t)$, with x in A . We will deduce from this an estimator of the cumulative rate $\Lambda(x, t)$ (see Theorem 4.21). An additional difficulty is related to the invariant measure of A , which we have to estimate.

We shall impose some assumptions about both the ergodicity of the Markov chain $(Z_n)_{n \geq 0}$ and the characteristics of the process. In the following, ν_n denotes the distribution of Z_n , for each integer n .

Assumption 4.1. *There exists a probability measure ν such that, for any initial distribution $\nu_0 = \delta_{\{x\}}$, $x \in E$,*

$$\lim_{n \rightarrow +\infty} \|\nu_n - \nu\|_{TV} = 0.$$

This assumption may be directly related to the transition kernel of the Markov chain $(Z_n)_{n \geq 0}$ (existence of a Foster-Lyapunov's function or Doeblin's condition for instance). We refer the interested reader to [22] for results about this kind of connection. This assumption leads to the following results.

Proposition 4.2. *We have the following statements:*

- (1) $(Z_n)_{n \geq 0}$ is ν -irreducible.
- (2) $(Z_n)_{n \geq 0}$ is positive Harris-recurrent and aperiodic.
- (3) ν is the unique invariant probability measure of $(Z_n)_{n \geq 0}$.

Proof.

- (1) Let $x \in E$ and $\nu_0 = \delta_{\{x\}}$. Then, from Assumption 4.1,

$$\lim_{n \rightarrow +\infty} \|Q^n(x, \cdot) - \nu\|_{TV} = 0.$$

Let us consider $A \in \mathcal{B}(E)$ such that $\nu(A) > 0$. Thus, $Q^n(x, A) > 0$ for n large enough. In the light of Proposition 4.2.1 of [22], the Markov chain $(Z_n)_{n \geq 0}$ is, therefore, ν -irreducible.

- (2) Recall that we have

$$\lim_{n \rightarrow +\infty} \|Q^n(x, \cdot) - \nu\|_{TV} = 0.$$

Thus, on the strength of Theorem 4.3.3 of [15], $(Z_n)_{n \geq 0}$ is positive Harris-recurrent and aperiodic.

- (3) A positive Harris-recurrent Markov chain admits a unique (up to a multiple constant) invariant measure (see for instance the introduction of [14]). \square

Denote by η_n the distribution of the couple (Z_n, S_{n+1}) , and by $\mu_z(\cdot)$ the conditional law of S_1 given $Z_0 = z$. By construction (3), the distribution of S_{n+1} given $Z_n = z$ is also given by $\mu_z(\cdot)$. The following result gives us the probability measure of (Z_n, S_{n+1}) according to $(\mu_z)_{z \in E}$ and ν_n .

Lemma 4.3. *For each integer n , η_n satisfies, for any $A \times \Gamma \in \mathcal{B}(E) \otimes \mathcal{B}(\mathbf{R}_+)$,*

$$\eta_n(A \times \Gamma) = \int_{A \times \Gamma} \mu_z(ds) \nu_n(dz).$$

Proof. This is the disintegration of the measure η_n according to its marginal distribution ν_n . \square

We shall see that the sequence $(\eta_n)_{n \geq 0}$ converges to the probability measure η given by,

$$\forall A \times \Gamma \in \mathcal{B}(E) \otimes \mathcal{B}(\mathbf{R}_+), \quad \eta(A \times \Gamma) = \int_{A \times \Gamma} \mu_z(ds) \nu(dz).$$

Lemma 4.4. *For any initial distribution $\nu_0 = \delta_{\{x\}}$, $x \in E$,*

$$\lim_{n \rightarrow +\infty} \|\eta_n - \eta\|_{TV} = 0.$$

Proof. Let g be a measurable function bounded by 1. According to Lemma 4.3,

$$\left| \int_{E \times \mathbf{R}_+} g(z, s) [\eta_n(dz \times ds) - \eta(dz \times ds)] \right| \leq \|h\|_\infty \|\nu_n - \nu\|_{TV},$$

on the strength of Fubini's theorem, where the function h is defined by,

$$\forall z \in E, \quad h(z) = \int_{\mathbf{R}_+} g(z, s) \mu_z(ds).$$

Furthermore, since $|h(z)| \leq 1$ for any z , we have

$$\left| \int_{E \times \mathbf{R}_+} g(z, s) (\eta_n(dz \times ds) - \eta(dz \times ds)) \right| \leq \|\nu_n - \nu\|_{TV},$$

Thus, by Assumption 4.1, $\|\eta_n - \eta\|_{TV}$ tends to 0. \square

As a consequence, the Markov chain $(Z_n, S_{n+1})_{n \geq 0}$ has similar properties of the Markov chain $(Z_n)_{n \geq 0}$ given in Proposition 4.2.

Proposition 4.5. *We have the following statements:*

- (1) $(Z_n, S_{n+1})_{n \geq 0}$ is η -irreducible.
- (2) $(Z_n, S_{n+1})_{n \geq 0}$ is positive Harris-recurrent and aperiodic.
- (3) η is the unique (up to a multiple constant) invariant probability measure of the Markov chain $(Z_n, S_{n+1})_{n \geq 0}$.

Proof. One may state this result from Lemma 4.4, with the arguments given in the proof of Proposition 4.2. \square

According to the previous discussion, we shall apply the ergodic theorem to the Markov chains $(Z_n)_{n \geq 0}$ and $(Z_n, S_{n+1})_{n \geq 0}$. Now, we impose some assumptions on the characteristics of the process.

Assumptions 4.6.

- (1) The jump rate λ is uniformly Lipschitz, that is,

$$\exists [\lambda]_{Lip} > 0, \forall \xi, \xi' \in E, \forall s \geq 0, |\lambda(\xi, s) - \lambda(\xi', s)| \leq [\lambda]_{Lip} d(\xi, \xi').$$
- (2) There exists a locally integrable function $M : \mathbf{R}_+ \rightarrow \mathbf{R}_+$ such that,

$$\forall \xi \in E, \forall s \geq 0, \lambda(\xi, s) \leq M(s).$$
- (3) The density f is continuous in time.
- (4) f is bounded.
- (5) The function t^* is continuous.

Under these assumptions, one states some intermediate results about λ , t^* and G .

Lemma 4.7. *Let $A \in \mathcal{B}(E)$ be a relatively compact set such that $\overline{A} \cap \partial E = \emptyset$. Thus,*

$$\inf_{\xi \in A} t^*(\xi) > 0.$$

In this case, one denotes $t^(A) = \inf_{\xi \in A} t^*(\xi)$. Furthermore,*

$$\forall t \geq 0, \inf_{\xi \in A} G(\xi, t) > 0.$$

Proof. By assumption, t^* is continuous. Moreover, for any $\xi \notin \partial E$, $t^*(\xi) > 0$. Since \overline{A} is a compact set such that $\overline{A} \cap \partial E = \emptyset$, we have

$$\inf_{\xi \in \overline{A}} t^*(\xi) > 0.$$

Hence,

$$\inf_{\xi \in A} t^*(\xi) \geq \inf_{\xi \in \overline{A}} t^*(\xi) > 0.$$

For any s , the function $\lambda(\cdot, s)$ is continuous because it is Lipschitz. Furthermore, $\lambda(\xi, \cdot)$ is bounded by M , not depending on ξ , and locally integrable. Recall that we have

$$G(\xi, t) = \exp \left(- \int_0^t \lambda(\xi, s) ds \right).$$

Thus, by Lebesgue's theorem of continuity under the integral sign, $G(\cdot, t)$ is continuous, and satisfies $G(\xi, t) > 0$ for any ξ . Hence,

$$\inf_{\xi \in A} G(\xi, t) \geq \inf_{\xi \in \bar{A}} G(\xi, t) > 0.$$

Both inequalities are, therefore, proved. \square

Let us introduce the following notation.

$$\mathcal{B}_\nu^+ = \{A \in \mathcal{B}(E) : A \text{ relatively compact, } \nu(A) > 0 \text{ and } \bar{A} \cap \partial E = \emptyset\}.$$

The following lemma states that, for any $A \in \mathcal{B}_\nu^+$, λ is bounded on $A \times [0, t^*(A)[$.

Lemma 4.8. *Let $A \in \mathcal{B}_\nu^+$, $\xi \in A$ and $0 \leq t < t^*(A)$. Thus,*

$$\lambda(\xi, t) \leq \frac{\|f\|_\infty}{\inf_{\xi \in A} G(\xi, t^*(A))}.$$

Proof. As f is bounded and $G(\xi, \cdot)$ is decreasing, we obtain by Lemma 4.7,

$$\lambda(\xi, t) \leq \frac{\|f\|_\infty}{G(\xi, t)} \leq \frac{\|f\|_\infty}{G(\xi, t^*(A))}.$$

This immediately yields to the expected result. \square

Let $A \in \mathcal{B}_\nu^+$. Let us consider for each integer n the continuous-time process $Y_n(A, \cdot)$, defined by,

$$(16) \quad \forall 0 \leq t < t^*(A), \quad Y_n(A, t) = \sum_{i=0}^{n-1} \mathbf{1}_{\{S_{i+1} \geq t\}} \mathbf{1}_{\{Z_i \in A\}}.$$

We shall state two lemmas about the asymptotic behavior of Y_n . Before, we focus our attention on the link between $G(z, \cdot)$ and $\mu_z(\cdot)$, for any $z \in A$.

Remark 4.9. *Let $A \in \mathcal{B}_\nu^+$ and $z \in A$. For any $0 \leq t < t^*(A)$, we have*

$$\begin{aligned} \mu_z([t, +\infty[) &= \mathbf{P}_{\nu_0}(S_1 \geq t | Z_0 = z) \\ &= \mathbf{P}_{\nu_0}(S_1 > t | Z_0 = z), \end{aligned}$$

because $t < t^*(A) \leq t^*(z)$. Thus,

$$\mu_z([t, +\infty[) = G(z, t).$$

Lemma 4.10. *Let $A \in \mathcal{B}_\nu^+$ and $x \in E$. Thus, for any $0 \leq t < t^*(A)$,*

$$\frac{Y_n(A, t)}{n} \longrightarrow \int_A G(z, t) \nu(dz) \quad \mathbf{P}_x\text{-a.s. as } n \rightarrow +\infty.$$

Furthermore, this limit is strictly positive.

Proof. In the light of Proposition 4.5, the Markov chain $(Z_n, S_{n+1})_{n \geq 0}$ is positive Harris-recurrent and admits η as its unique invariant probability measure. Thus, using the ergodic theorem (see for instance Theorem 17.1.7 of [22]), we have

$$\frac{Y_n(A, t)}{n} \rightarrow \int_A \mu_z([t, +\infty[) \nu(dz) \quad \mathbf{P}_x\text{-a.s. as } n \rightarrow +\infty.$$

Furthermore, for any $z \in A$, as $t < t^*(A)$ and according to Remark 4.9,

$$\mu_z([t, +\infty[) = G(z, t).$$

It is a strictly positive number because $\nu(A) > 0$ and $\inf_{\xi \in A} G(\xi, t) > 0$ by Lemma 4.7. \square

Let $A \in \mathcal{B}_\nu^+$ and $0 \leq t < t^*(A)$. Let us introduce the generalized inverse $Y_n(A, t)^+$ of $Y_n(A, t)$ by

$$(17) \quad Y_n(A, t)^+ = \begin{cases} 0 & \text{if } Y_n(A, t) = 0, \\ \frac{1}{Y_n(A, t)} & \text{else.} \end{cases}$$

Lemma 4.11. *Let $A \in \mathcal{B}_\nu^+$, $0 \leq t < t^*(A)$ and $x \in E$. Thus, for all integers n ,*

$$(18) \quad Y_n(A, t)^+ \leq 1 \text{ } \mathbf{P}_x\text{-a.s.}$$

and, as n goes to infinity,

$$(19) \quad Y_n(A, t)^+ \longrightarrow 0 \text{ } \mathbf{P}_x\text{-a.s.},$$

$$(20) \quad \mathbf{1}_{\{Y_n(A, t)=0\}} \longrightarrow 0 \text{ } \mathbf{P}_x\text{-a.s.},$$

$$\int_0^t \mathbf{1}_{\{Y_n(A, s)=0\}} ds \longrightarrow 0 \text{ } \mathbf{P}_x\text{-a.s.}$$

Proof. $Y_n(A, t)^+$ is almost surely bounded by 1, since $Y_n(A, t)$ takes its values on the integers. One immediately obtains the limits (19) and (20), because $Y_n(A, t)/n$ almost surely admits a strictly positive limit by virtue of Lemma 4.10. Finally,

$$\limsup_{n \rightarrow +\infty} \int_0^t \mathbf{1}_{\{Y_n(A, s)=0\}} ds \leq \int_0^t \limsup_{n \rightarrow +\infty} \mathbf{1}_{\{Y_n(A, s)=0\}} ds = 0,$$

by (20). \square

In the following proposition, we shall apply the ergodic theorem in order to define, for any $A \in \mathcal{B}_\nu^+$, the function $l(A, \cdot)$ which is an approximation of the jump rate $\lambda(\xi, \cdot)$, for $\xi \in A$ (see Lemma 4.18). We also state the continuity of this function.

Proposition 4.12. *Let $A \in \mathcal{B}_\nu^+$, $0 \leq t < t^*(A)$ and $x \in E$. Thus, when n goes to infinity,*

$$(21) \quad Y_n(A, t)^+ \sum_{i=0}^{n-1} \lambda(Z_i, t) \mathbf{1}_{\{Z_i \in A\}} \mathbf{1}_{\{S_{i+1} \geq t\}} \longrightarrow l(A, t) = \frac{\int_A f(z, t) \nu(dz)}{\int_A G(z, t) \nu(dz)} \text{ } \mathbf{P}_x\text{-a.s.}$$

The function $l(A, \cdot)$ is continuous on $[0, t^*(A)[$. We especially have

$$(22) \quad K_t(A) = \sup_{0 \leq s \leq t} |l(A, s)| < +\infty.$$

Proof. The ergodic theorem (Theorem 17.1.7 of [22]) applied to $(Z_n, S_{n+1})_{n \geq 0}$ leads to

$$\begin{aligned} \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{i=0}^{n-1} \lambda(Z_i, t) \mathbf{1}_{\{Z_i \in A\}} \mathbf{1}_{\{S_{i+1} \geq t\}} &= \int_A \lambda(z, t) \mu_z([t, +\infty[) \nu(dz) \text{ } \mathbf{P}_x\text{-a.s.} \\ &= \int_A \lambda(z, t) G(z, t) \nu(dz) \text{ } \mathbf{P}_x\text{-a.s.}, \end{aligned}$$

because $t < t^*(A)$. Notice that $f(z, t) = \lambda(z, t)G(z, t)$. Thus,

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{i=0}^{n-1} \lambda(Z_i, t) \mathbf{1}_{\{Z_i \in A\}} \mathbf{1}_{\{S_{i+1} \geq t\}} = \int_A f(z, t) \nu(dz) \quad \mathbf{P}_x\text{-a.s.}$$

Furthermore, in the light of Lemma 4.10,

$$\lim_{n \rightarrow +\infty} \frac{1}{n} Y_n(A, t) = \int_A G(z, t) \nu(dz) \quad \mathbf{P}_x\text{-a.s.}$$

Finally, doing the ratio of these two limits leads to the expected convergence. On the strength of Lebesgue's theorem of continuity under the integral sign, $l(A, \cdot)$ is a continuous function, since f and G are continuous in time and bounded. \square

Having established the asymptotic behavior of Y_n , we focus on continuous-time martingales. In particular, thanks to Lenglart's inequality for continuous-time martingales, we shall estimate, for any $A \in \mathcal{B}_\nu^+$, the functions $l(A, \cdot)$ and $L(A, \cdot)$, where $L(A, \cdot)$ is given by,

$$(23) \quad \forall 0 \leq t < t^*(A), \quad L(A, t) = \int_0^t l(A, s) ds.$$

Theorem 4.13. *Let $A \in \mathcal{B}_\nu^+$. For each integer n , the process $M_n(A, \cdot)$ defined by,*

$$(24) \quad \forall 0 \leq t < t^*(A), \quad M_n(A, t) = \sum_{i=0}^{n-1} M^{i+1}(t) \mathbf{1}_{\{Z_i \in A\}},$$

is a continuous-time martingale in the filtration $(\mathcal{G}_n \vee \bigvee_{i=0}^{n-1} \mathcal{F}_t^{i+1})_{0 \leq t < t^(A)}$.*

Proof. Let $0 \leq s < t < t^*(A)$. Then,

$$\mathbf{E}_{\nu_0} [M_n(A, t) | \mathcal{G}_n \vee \bigvee_{i=0}^{n-1} \mathcal{F}_s^{i+1}] = \sum_{i=0}^{n-1} \mathbf{E}_{\nu_0} [M^{i+1}(t) \mathbf{1}_{\{Z_i \in A\}} | \mathcal{G}_n \vee \bigvee_{j=0}^{n-1} \mathcal{F}_s^{j+1}].$$

Moreover, on the strength of Proposition 2.1, we have

$$\bigvee_{j \neq i+1} \mathcal{F}_s^j \perp_{\mathcal{G}_n \vee \mathcal{F}_s^{i+1}} \mathcal{F}_t^{i+1} \quad \text{and} \quad \mathcal{F}_t^{i+1} \perp_{\sigma(Z_i) \vee \mathcal{F}_s^{i+1}} \mathcal{G}_n.$$

Therefore, since $\sigma(Z_i)$ is a sub- σ -field of \mathcal{G}_n , we have by Corollary 6.8 of [17],

$$\bigvee_{j \neq i+1} \mathcal{F}_s^j \perp_{\mathcal{G}_n \vee \mathcal{F}_s^{i+1}} \mathcal{F}_t^{i+1} \vee \sigma(Z_i) \quad \text{and} \quad \sigma(Z_i) \vee \mathcal{F}_t^{i+1} \perp_{\sigma(Z_i) \vee \mathcal{F}_s^{i+1}} \mathcal{G}_n.$$

Thus, as $M^{i+1}(t) \mathbf{1}_{\{Z_i \in A\}}$ is $\sigma(Z_i) \vee \mathcal{F}_t^{i+1}$ -measurable,

$$\begin{aligned} \mathbf{E}_{\nu_0} [M_n(A, t) | \mathcal{G}_n \vee \bigvee_{i=0}^{n-1} \mathcal{F}_s^{i+1}] &= \sum_{i=0}^{n-1} \mathbf{E}_{\nu_0} [M^{i+1}(t) \mathbf{1}_{\{Z_i \in A\}} | \mathcal{G}_n \vee \mathcal{F}_s^{i+1}] \\ &= \sum_{i=0}^{n-1} \mathbf{E}_{\nu_0} [M^{i+1}(t) \mathbf{1}_{\{Z_i \in A\}} | \sigma(Z_i) \vee \mathcal{F}_s^{i+1}]. \end{aligned}$$

Furthermore, with Lemma 2.2,

$$\mathbf{E}_{\nu_0} [M^{i+1}(t) \mathbf{1}_{\{Z_i \in A\}} | \sigma(Z_i) \vee \mathcal{F}_s^{i+1}] = M^{i+1}(s) \mathbf{1}_{\{Z_i \in A\}}.$$

Thus,

$$\mathbf{E}_{\nu_0} [M_n(A, t) | \mathcal{G}_n \vee \bigvee_{i=0}^{n-1} \mathcal{F}_s^{i+1}] = \sum_{i=0}^{n-1} M^{i+1}(s) \mathbf{1}_{\{Z_i \in A\}}.$$

By (24), this ensures that $M_n(A, \cdot)$ is a martingale. \square

For any $A \in \mathcal{B}_\nu^+$, let us introduce the counting process $N_n(A, \cdot)$, defined for any $0 \leq t < t^*(A)$ by

$$(25) \quad N_n(A, t) = \sum_{i=0}^{n-1} \mathbf{1}_{\{S_{i+1} \leq t\}} \mathbf{1}_{\{Z_i \in A\}}.$$

Lemma 4.14. *Let $A \in \mathcal{B}_\nu^+$. For all integers n , the process given for all $0 \leq s < t^*(A)$ by*

$$(26) \quad \widetilde{M}_n(A, s) = \int_0^s Y_n(A, u)^+ dM_n(A, u),$$

is a martingale whose predictable variation process $\langle \widetilde{M}_n(A) \rangle$ satisfies for any $x \in E$,

$$\forall 0 \leq s < t^*(A), \langle \widetilde{M}_n(A) \rangle(s) \rightarrow 0 \quad \mathbf{P}_x\text{-a.s. as } n \rightarrow +\infty.$$

Proof. By (9), (24) and (25), for any $0 \leq t < t^*(A)$, one may differently write $M_n(A, t)$,

$$(27) \quad M_n(A, t) = N_n(A, t) - \int_0^t \sum_{i=0}^{n-1} \mathbf{1}_{\{S_{i+1} \geq u\}} \mathbf{1}_{\{Z_i \in A\}} \lambda(Z_i, u) du.$$

In the light of Theorem 4.13, this is a continuous-time martingale. As a consequence, the process $\mathcal{A}_n(A, \cdot)$ given by,

$$\forall 0 \leq s < t^*(A), \mathcal{A}_n(A, s) = \int_0^s \sum_{i=0}^{n-1} \mathbf{1}_{\{S_{i+1} \geq u\}} \mathbf{1}_{\{Z_i \in A\}} \lambda(Z_i, u) du,$$

is the compensator of the counting process $N_n(A, \cdot)$. In order to prove that $\widetilde{M}_n(A, \cdot)$ is a martingale, one may only state that

$$\mathbf{E}_{\nu_0} \left[\int_0^t \left(Y_n(A, s)^+ \right)^2 d\mathcal{A}_n(A, s) \right] < +\infty.$$

Recall that λ is bounded on the set $A \times [0, t^*(A)[$ on the strength of Lemma 4.8. C denotes an upper bound of λ on this set. Consequently,

$$\begin{aligned} \int_0^t \left(Y_n(A, s)^+ \right)^2 d\mathcal{A}_n(A, s) &= \sum_{i=0}^{n-1} \int_0^t \left(Y_n(A, s)^+ \right)^2 \mathbf{1}_{\{S_{i+1} \geq u\}} \mathbf{1}_{\{Z_i \in A\}} \lambda(Z_i, s) ds \\ &\leq \int_0^t \left(Y_n(A, s)^+ \right)^2 C \sum_{i=0}^{n-1} \mathbf{1}_{\{S_{i+1} \geq s\}} \mathbf{1}_{\{Z_i \in A\}} ds, \end{aligned}$$

Furthermore, by (16) and (17),

$$(28) \quad \int_0^t \left(Y_n(A, s)^+ \right)^2 d\mathcal{A}_n(A, s) \leq C \int_0^t Y_n(A, s)^+ ds.$$

As $Y_n(A, \cdot)^+$ is bounded by 1 by (18), we have

$$\int_0^t Y_n(A, s)^+ d\mathcal{A}_n(A, s) \leq Ct.$$

Hence,

$$\mathbf{E}_{\nu_0} \left[\int_0^t \left(Y_n(A, s)^+ \right)^2 d\mathcal{A}_n(A, s) \right] < +\infty.$$

This states that $\widetilde{M}_n(A, \cdot)$ is a martingale. The predictable variation of $\widetilde{M}_n(A, \cdot)$ is given by,

$$\forall 0 \leq s < t^*(A), \quad \langle \widetilde{M}_n(A) \rangle(s) = \int_0^s \left(Y_n(A, u)^+ \right)^2 d\mathcal{A}_n(A, u).$$

Therefore, by (28),

$$\forall 0 \leq s < t^*(A), \quad \langle \widetilde{M}_n(A) \rangle(s) \leq C \int_0^s Y_n(A, u)^+ du.$$

According to Lemma 4.11, $Y_n(A, \cdot)^+$ is bounded by 1 and for any u , $Y_n(A, u)^+$ almost surely tends to 0. Thus, by Lebesgue's dominated convergence theorem,

$$\langle \widetilde{M}_n(A) \rangle(s) \longrightarrow 0 \text{ } \mathbf{P}_x\text{-a.s. when } n \rightarrow +\infty.$$

This achieves the proof. \square

We stated this lemma in order to apply Lengart's inequality to the martingale $\widetilde{M}_n(A, \cdot)$. A reference about this inequality for continuous-time martingales may be found in [5], II.5.2.1. Lengart's inequality.

Remark 4.15. *In the light of Lengart's inequality, the previous lemma directly induces that for any $A \in \mathcal{B}_\nu^+$, $0 \leq t < t^*(A)$, and $x \in E$,*

$$\sup_{0 \leq s \leq t} \left| \widetilde{M}_n(A, s) \right| \xrightarrow{\mathbf{P}_x} 0.$$

Let $A \in \mathcal{B}_\nu^+$. We propose $\widehat{L}_n(A, \cdot)$ as an estimator of the function $L(A, \cdot)$ defined by (23). It is given by,

$$(29) \quad \forall 0 \leq t < t^*(A), \quad \widehat{L}_n(A, t) = \int_0^t Y_n(A, s)^+ dN_n(A, s).$$

$\widehat{L}_n(A, \cdot)$ is a Nelson-Aalen type estimator of $L(A, \cdot)$. The following results are related to its asymptotic behavior. We shall see that smoothing this estimator provides an estimator of $l(A, \cdot)$. Before, let us introduce this notation, for each $n \geq 0$,

$$(30) \quad \forall 0 \leq t < t^*(A), \quad L_n^*(A, t) = \int_0^t l(A, s) \mathbf{1}_{\{Y_n(A, s) > 0\}} ds.$$

We recall that $Y_n(A, \cdot)$ and its generalized inverse have already been defined by (16) and (17).

Proposition 4.16. *Let $A \in \mathcal{B}_\nu^+$, $0 < t < t^*(A)$ and $x \in E$. Then,*

$$\sup_{0 \leq s \leq t} \left| \widehat{L}_n(A, s) - L_n^*(A, s) \right| \xrightarrow{\mathbf{P}_x} 0,$$

when n goes to infinity.

Proof. The definition of $\widetilde{M}_n(A, s)$ (26), the expression of $M_n(A, s)$ (27) and the definition of $\widehat{L}_n(A, s)$ (29) yield to

$$\widetilde{M}_n(A, s) = \widehat{L}_n(A, s) - \int_0^s Y_n(A, u)^+ \lambda(Z_i, u) \mathbf{1}_{\{Z_i \in A\}} \mathbf{1}_{\{S_{i+1} \geq u\}} du.$$

Thus, $\widetilde{M}_n(A, s)$ may be written in the following way.

$$\widetilde{M}_n(A, s) = \widehat{L}_n(A, s) - L_n^*(A, s) - a_n(s),$$

where the extra-term $a_n(s)$ is given by

$$a_n(s) = \int_0^s Y_n(A, u)^+ \sum_{i=0}^{n-1} [\lambda(Z_i, u) - l(A, u)] \mathbf{1}_{\{Z_i \in A\}} \mathbf{1}_{\{S_{i+1} \geq u\}} du.$$

Therefore,

$$\sup_{0 \leq s \leq t} \left| \widehat{L}_n(A, s) - L_n^*(A, s) \right| \leq \sup_{0 \leq s \leq t} |\widetilde{M}_n(A, s)| + \sup_{0 \leq s \leq t} |a_n(s)|.$$

We have already stated in Remark 4.15 that $\sup_{0 \leq s \leq t} |\widetilde{M}_n(A, s)|$ tends in probability to 0. As a consequence, we only need to study the limit of $\sup_{0 \leq s \leq t} |a_n(s)|$. With (16), we have

$$|a_n(s)| \leq \int_0^t \mathbf{1}_{\{Y_n(A, u) > 0\}} \left| Y_n(A, u)^+ \sum_{i=0}^{n-1} \lambda(Z_i, u) \mathbf{1}_{\{Z_i \in A\}} \mathbf{1}_{\{S_{i+1} \geq u\}} - l(A, u) \right| du.$$

The integrated function almost surely converges to 0 on the strength of Proposition 4.12. Furthermore, in the light of Lemma 4.8, there exists a real number $C > 0$, which is an upper bound of λ on $A \times [0, t^*(A)]$. Hence, for $u \leq t$,

$$\begin{aligned} \left| Y_n(A, u)^+ \sum_{i=0}^{n-1} \lambda(Z_i, u) \mathbf{1}_{\{Z_i \in A\}} \mathbf{1}_{\{S_{i+1} \geq u\}} - l(A, u) \right| &\leq |l(A, u)| + C \\ &\leq K_t(A) + C, \end{aligned}$$

where $K_t(A)$ has already been defined by (22). As a consequence, we apply Lebesgue's dominated convergence theorem, and we obtain

$$\lim_{n \rightarrow +\infty} \sup_{0 \leq s \leq t} |a_n(s)| = 0 \quad \mathbf{P}_x\text{-a.s.},$$

showing the result. \square

Theorem 4.17. *Let $A \in \mathcal{B}_\nu^+$, $0 < t < t^*(A)$ and $x \in E$. Then,*

$$\sup_{0 \leq s \leq t} |\widehat{L}_n(A, s) - L(A, s)| \xrightarrow{\mathbf{P}_x} 0 \text{ as } n \rightarrow +\infty.$$

Proof. By the triangle inequality, we have

$$|\widehat{L}_n(A, s) - L(A, s)| \leq |\widehat{L}_n(A, s) - L_n^*(A, s)| + |L_n^*(A, s) - L(A, s)|.$$

In the light of Proposition 4.16, we only need to show that $\sup_{0 \leq s \leq t} |L_n^*(A, s) - L(A, s)|$ tends in probability to 0. By (22), (23) and (30),

$$\sup_{0 \leq s \leq t} |L_n^*(A, s) - L(A, s)| \leq K_t(A) \int_0^t \mathbf{1}_{\{Y_n(A, r) = 0\}} dr,$$

and the bound almost surely converges to 0 by Lemma 4.11. \square

Let $A \in \mathcal{B}_\nu^+$. We have provided an estimator of the function $L(A, \cdot)$ on the interval $[0, t^*(A)[$ (see Theorem 4.17). We shall prove that $L(A, \cdot)$ and the cumulative rate $\Lambda(\xi, \cdot)$, defined by (5), are close for any $\xi \in A$ if A is small enough. In the same way, we state that $l(A, \cdot)$, given by (21), and the jump rate $\lambda(\xi, \cdot)$ are close for any $\xi \in A$.

Lemma 4.18. *Let $A \in \mathcal{B}_\nu^+$, $z \in A$ and $0 \leq s < t^*(A)$,*

$$\begin{aligned} |\lambda(z, s) - l(A, s)| &\leq [\lambda]_{Lip} \text{diam } A, \\ |\Lambda(z, s) - L(A, s)| &\leq s [\lambda]_{Lip} \text{diam } A. \end{aligned}$$

Proof. First, we have

$$\begin{aligned} |\lambda(z, s) - l(A, s)| &= \left| \lambda(z, s) - \frac{\int_A f(\xi, s) \nu(d\xi)}{\int_A G(\xi, s) \nu(d\xi)} \right| \\ &\leq \frac{1}{\int_A G(\xi, s) \nu(d\xi)} \int_A |\lambda(z, s) - \lambda(\xi, s)| G(\xi, s) \nu(d\xi) \\ &\leq [\lambda]_{Lip} \text{diam } A. \end{aligned}$$

Moreover, integrating the previous result leads to $|\Lambda(z, s) - L(A, s)| \leq s [\lambda]_{Lip} \text{diam } A$. \square

Notice that we can define and estimate the function $L(A, \cdot)$ only when $\nu(A)$ is strictly positive. Therefore, we need to estimate the indicator function $\mathbf{1}_{\{\nu(A) > 0\}}$. For this, we estimate the quantity $\nu(A)$ by its empirical version,

$$\hat{\nu}_n(A) = \frac{1}{n} \sum_{i=0}^{n-1} \mathbf{1}_{\{Z_i \in A\}}.$$

Lemma 4.19. *By virtue of the ergodic theorem, $\hat{\nu}_n(A) \rightarrow \nu(A)$ almost surely. Furthermore,*

$$\mathbf{1}_{\{\hat{\nu}_n(A) > n^{-1/2}\}} \rightarrow \mathbf{1}_{\{\nu(A) > 0\}} \quad \mathbf{P}_x\text{-a.s. when } n \rightarrow +\infty,$$

for any $x \in E$.

Proof. Let us distinguish the cases $\nu(A) > 0$ and $\nu(A) = 0$.

- (i) If $\nu(A) > 0$, $\hat{\nu}_n(A)$ has an almost sure limit which is strictly positive. Thus, for n large enough, $\hat{\nu}_n(A) > n^{-1/2}$ almost surely.
- (ii) If $\nu(A) = 0$, the number of visits in A of the Markov chain $(Z_n)_{n \geq 0}$ is almost surely finite, because this Markov chain is Harris-recurrent. In this case, the sum $\sum_{i=0}^{n-1} \mathbf{1}_{\{Z_i \in A\}}$ almost surely converges to a finite sum. \square

In the following remark, we focus on the asymptotic behavior of $\sup_{0 \leq s \leq t} \hat{L}_n(A, s)$ when $\nu(A) = 0$. Indeed, we stated that the convergence of the estimator holds only when $\nu(A) > 0$.

Remark 4.20. *Let $A \in \mathcal{B}(E)$ such that $\bar{A} \cap \partial E = \emptyset$ and $\nu(A) = 0$. Let $0 \leq s \leq t < t^*(A)$ and $x \in E$. By (29), we have*

$$\hat{L}_n(A, s) \leq \sum_{i=0}^{n-1} Y_n(A, S_{i+1})^+ \mathbf{1}_{\{Z_i \in A\}} \quad \mathbf{P}_x\text{-a.s.}$$

Together with (18), we obtain

$$\sup_{0 \leq s \leq t} \widehat{L}_n(A, s) \leq \sum_{i=0}^{n-1} \mathbf{1}_{\{Z_i \in A\}} \mathbf{P}_x\text{-a.s.}$$

In addition, the number of visits in A is almost surely finite since $\nu(A) = 0$. Thus, $\sup_{0 \leq s \leq t} \widehat{L}_n(A, s)$ almost surely tends to a finite sum.

For estimating Λ on $\mathcal{K} \times [0, t]$ where \mathcal{K} is a compact subset of E , one considers a thin enough partition (A_k) of \mathcal{K} and one estimates $\Lambda(\xi, s)$ by $\widehat{L}_n(A_j, s)$, for $\xi \in A_j$.

Theorem 4.21. *Let \mathcal{K} be a compact subset of E and $\xi \in E$. For any $\varepsilon, \eta > 0$, there exist an integer N and a finite partition $P = (A_k)$ of \mathcal{K} , such that for any $n \geq N$, for any $0 < t < \min_k t^*(A_k)$,*

$$\mathbf{P}_\xi \left(\sup_{x \in \mathcal{K}} \sup_{0 \leq s \leq t} \left| \sum_{k=1}^{|P|} \widehat{L}_n(A_k, s) \mathbf{1}_{\{\widehat{\nu}_n(A_k) > \frac{1}{\sqrt{n}}\}} \mathbf{1}_{\{x \in A_k\}} - \Lambda(x, s) \mathbf{1}_{\{x \in \mathcal{K}'\}} \right| > \eta \right) < \varepsilon,$$

where \mathcal{K}' is defined by

$$\mathcal{K}' = \bigcup_{\nu(A_k) > 0} A_k.$$

Proof. Let us fix s and x . We have

$$\begin{aligned} & \sum_{k=1}^{|P|} \widehat{L}_n(A_k, s) \mathbf{1}_{\{\widehat{\nu}_n(A_k) > \frac{1}{\sqrt{n}}\}} \mathbf{1}_{\{x \in A_k\}} - \Lambda(x, s) \mathbf{1}_{\{x \in \mathcal{K}'\}} \\ &= \sum_{k=1}^{|P|} \widehat{L}_n(A_k, s) \mathbf{1}_{\{\widehat{\nu}_n(A_k) > \frac{1}{\sqrt{n}}\}} \mathbf{1}_{\{x \in A_k\}} - \sum_{k=1}^{|P|} \Lambda(x, s) \mathbf{1}_{\{x \in A_k\}} \mathbf{1}_{\{\nu(A_k) > 0\}} \\ &= \sum_{k=1}^{|P|} \widehat{L}_n(A_k, s) \mathbf{1}_{\{\widehat{\nu}_n(A_k) > \frac{1}{\sqrt{n}}\}} \mathbf{1}_{\{x \in A_k\}} \mathbf{1}_{\{\nu(A_k) = 0\}} \\ &+ \sum_{k=1}^{|P|} \left(\widehat{L}_n(A_k, s) \mathbf{1}_{\{\widehat{\nu}_n(A_k) > \frac{1}{\sqrt{n}}\}} - \Lambda(x, s) \right) \mathbf{1}_{\{x \in A_k\}} \mathbf{1}_{\{\nu(A_k) > 0\}}. \end{aligned}$$

Let us denote by l the integer between 1 and $|P|$ such that x is in A_l . We distinguish the two cases $\nu(A_l) > 0$ and $\nu(A_l) = 0$. By assumption on \mathcal{K} , A_l is a relatively compact set such that $\overline{A_l} \cap \partial E = \emptyset$.

- (i) If $\nu(A_l) = 0$, $\widehat{L}_n(A_l, s)$ almost surely and uniformly converges to a finite sum, by Remark 4.20. Thus, by Lemma 4.19,

$$\sup_{0 \leq s \leq t} \widehat{L}_n(A_l, s) \mathbf{1}_{\{\widehat{\nu}_n(A_l) > \frac{1}{\sqrt{n}}\}} \xrightarrow{\mathbf{P}_x} 0.$$

(ii) If $\nu(A_l) > 0$. We have, by the triangle inequality,

$$\begin{aligned} |\widehat{L}_n(A_l, s) \mathbf{1}_{\{\widehat{\nu}_n(A_l) > \frac{1}{\sqrt{n}}\}} - \Lambda(x, s)| &\leq |\widehat{L}_n(A_l, s) \mathbf{1}_{\{\widehat{\nu}_n(A_l) > n^{-1/2}\}} - 1| \\ &\quad + |\widehat{L}_n(A_l, s) - L(A_l, s)| \\ &\quad + |L(A_l, s) - \Lambda(x, s)|. \end{aligned}$$

- (a) The first term uniformly tends to 0 in probability, since $|\mathbf{1}_{\{\widehat{\nu}_n(A_l) > n^{-1/2}\}} - 1|$ almost surely tends to 0 by Lemma 4.19, and because $\widehat{L}_n(A_l, \cdot)$ uniformly converges in probability to $L(A_l, \cdot)$, according to Theorem 4.17.
- (b) The second term uniformly tends to 0 in probability on the strength of Theorem 4.17.
- (c) The third term is bounded by $t[\lambda]_{Lip} \max_k \text{diam } A_k$, according to Lemma 4.18, which does not depend on x and is arbitrarily small.

Since the sum is finite, this achieves the proof. \square

4.2. Estimation of λ . In this part, we focus on smoothing the estimator $\widehat{L}_n(A, \cdot)$ of $L(A, \cdot)$ in order to provide a consistent estimator of $l(A, \cdot)$ (see Proposition 4.23), and, therefore, of λ (see Theorem 4.24). Indeed, we shall state another corollary of Proposition 4.16, which deals with the smoothing of $\widehat{L}_n(A, \cdot)$ by some kernel methods.

Let K be a continuous kernel with support $[-1, 1]$. Let us introduce the following notations. For any real number $b > 0$ and $0 < t < t^*(A)$, we denote,

$$\forall 0 \leq u \leq t, \widehat{l}_{n,b,t}(A, u) = \frac{1}{b} \int_0^t K\left(\frac{u-s}{b}\right) d\widehat{L}_n(A, s).$$

This is an estimator of $l(A, \cdot)$ within the interval $[0, t]$. Furthermore, we denote also,

$$\forall 0 \leq u \leq t, l_{n,b,t}^*(A, u) = \frac{1}{b} \int_0^t K\left(\frac{u-s}{b}\right) dL_n^*(A, s).$$

We first prove the following lemma.

Lemma 4.22. *Let $A \in \mathcal{B}_\nu^+$, $b > 0$ and $0 < r < t < t^*(A)$.*

$$\sup_{0 \leq s \leq r} |\widehat{l}_{n,b,t}(A, s) - l_{n,b,t}^*(A, s)| \leq \frac{2}{b} V(K) \sup_{0 \leq s \leq t} |\widehat{L}_n(A, s) - L_n^*(A, s)|.$$

Proof. Let us denote,

$$\forall 0 \leq s \leq t, g(s) = \widehat{L}_n(A, s) - L_n^*(A, s).$$

Thus, by an integration by parts and since $g(0) = 0$,

$$K((s-t)/b)g(t) = \int_0^t K((s-u)/b) dg(u) - \int_0^t g(u-) dK((s-u)/b).$$

We deduce from the triangle inequality,

$$b \sup_{0 \leq s \leq r} |\widehat{l}_{n,b,t}(A, s) - l_{n,b,t}^*(A, s)| \leq \sup_{0 \leq s \leq r} |K((s-t)/b)g(t)| + \sup_{0 \leq s \leq r} \int_0^t g(u-) dK((s-u)/b).$$

For any $v \notin [-1, 1]$, $K(v) = 0$. This induces that

$$\begin{aligned} |K((s-t)/b)g(t)| &\leq \sup_{0 \leq u \leq t} |g(u)|K((s-t)/b) \\ &= \sup_{0 \leq u \leq t} |g(u)| \left\{ K((s-t)/b) - K(v) \right\} \\ &\leq \sup_{0 \leq u \leq t} |g(u)| V(K). \end{aligned}$$

Furthermore,

$$\begin{aligned} \int_0^t g(u-) dK((s-u)/b) &= \lim_{\max\{s_{i+1}-s_i\} \rightarrow 0} \sum_{i=1}^p g(s_i) \left\{ K((s-s_{i+1})/b) - K((s-s_i)/b) \right\} \\ &\leq \sup_{0 \leq u \leq t} |g(u)| V(K), \end{aligned}$$

giving the proof. \square

Proposition 4.23. *Let $A \in \mathcal{B}_v^+$ and $0 < r_1 < r_2 < t < t^*(A)$. There exists a sequence $(\beta_n)_{n \geq 0}$, which almost surely tends to 0, such that*

$$\sup_{r_1 \leq s \leq r_2} |\widehat{l}_{n,\beta_n,t}(A, s) - l(A, s)| \xrightarrow{\mathbf{P}^x} 0 \text{ when } n \rightarrow +\infty,$$

for any $x \in E$.

Proof. By the triangle inequality, we have for any sequence $(b_n)_{n \geq 0}$,

$$(31) \quad |\widehat{l}_{n,b_n,t}(A, s) - l(A, s)| \leq |\widehat{l}_{n,b_n,t}(A, s) - l_{n,b_n,t}^*(A, s)| + |l_{n,b_n,t}^*(A, s) - l(A, s)|.$$

We consider the sequence $(b_n)_{n \geq 0}$ defined by,

$$\forall n \geq 0, b_n = \sqrt{\sup_{0 \leq s \leq t} |\widehat{L}_n(A, s) - L_n^*(A, s)|}.$$

On the strength of Proposition 4.16, the sequence $(b_n)_{n \geq 0}$ tends in probability to 0, and we have

$$\sup_{0 \leq s \leq t} |\widehat{L}_n(A, s) - L_n^*(A, s)| \stackrel{\mathbf{P}^x}{=} o(b_n) \text{ when } n \rightarrow +\infty.$$

Finally, according to Lemma 4.22,

$$(32) \quad \sup_{r_1 \leq s \leq r_2} |\widehat{l}_{n,b_n,t}(A, s) - l_{n,b_n,t}^*(A, s)| \xrightarrow{\mathbf{P}^x} 0 \text{ when } n \rightarrow +\infty.$$

Now, we focus on the asymptotic behavior of $|l_{n,b_n,t}^*(A, s) - l(A, s)|$. By the change of variable $s - b_nv = u$, we have

$$l_{n,b_n,t}^*(A, s) = \int_{(s-t)/b_n}^{s/b_n} K(v) \mathbf{1}_{\{Y_n(A, s-b_nv) > 0\}} l(A, s - b_nv) dv.$$

As $(b_n)_{n \geq 0}$ tends in probability to 0, there exists a subsequence

$$(\beta_n)_{n \geq 0} = (b_{\alpha(n)})_{n \geq 0},$$

which almost surely converges to 0. Let

$$(33) \quad \Omega_1 = \{\omega \in \Omega : \beta_n(\omega) \rightarrow 0 \text{ and } \mathbf{1}_{\{Y_n(A, r_2, \omega) = 0\}} \rightarrow 0 \text{ when } n \rightarrow +\infty\}.$$

According to foregoing and by (20), $\mathbf{P}_x(\Omega_1) = 1$. Let $\omega \in \Omega_1$ and N such that, for any $n \geq N$,

$$r_1/\beta_n(\omega) \geq 1 \quad \text{and} \quad (r_2 - t)/\beta_n(\omega) \leq -1.$$

For any $n \geq N$, we have

$$\frac{s}{\beta_n(\omega)} \geq \frac{r_1}{\beta_n(\omega)} \geq 1 \quad \text{and} \quad \frac{s-t}{\beta_n(\omega)} \leq \frac{r_2-t}{\beta_n(\omega)} \leq -1.$$

Thus,

$$l_{n,\beta_n(\omega),t}^*(A, s, \omega) = \int_{-1}^1 K(v) \mathbf{1}_{\{Y_n(A, s-\beta_n(\omega)v, \omega) > 0\}} l(A, s - \beta_n(\omega)v) dv,$$

because the support of K is $[-1, 1]$. Hence,

$$\begin{aligned} & |l_{n,\beta_n(\omega),t}^*(A, s, \omega) - l(A, s)| \\ & \leq \left| \int_{-1}^1 K(u) \left(l(A, s - \beta_n(\omega)u) - l(A, s) \right) du \right| \\ (34) \quad & + \left| \int_{-1}^1 K(v) \left(\mathbf{1}_{\{Y_n(A, s-\beta_n(\omega)v, \omega) > 0\}} - 1 \right) l(A, s - \beta_n(\omega)v) dv \right|. \end{aligned}$$

We shall prove that the terms

$$\sup_{r_1 \leq s \leq r_2} \left| \int_{-1}^1 K(u) \left(l(A, s - \beta_n(\omega)u) - l(A, s) \right) du \right|$$

and

$$\sup_{r_1 \leq s \leq r_2} \left| \int_{-1}^1 K(v) \left(\mathbf{1}_{\{Y_n(A, s-\beta_n(\omega)v, \omega) > 0\}} - 1 \right) l(A, s - \beta_n(\omega)v) dv \right|$$

converge to 0 for any $\omega \in \Omega_1$. We shall begin by the second term.

(i) Recall that K and $l(A, \cdot)$ are two continuous functions. Let C defined by

$$C = K_t(A) \sup_{-1 \leq v \leq 1} K(v),$$

where $K_t(A)$ has already been defined by (22). Then, we have

$$\begin{aligned} & \left| \int_{-1}^1 K(v) \left(\mathbf{1}_{\{Y_n(A, s-\beta_n(\omega)v, \omega) > 0\}} - 1 \right) l(A, s - \beta_n(\omega)v) dv \right| \\ & \leq C \int_{-1}^1 \left| \mathbf{1}_{\{Y_n(A, s-\beta_n(\omega)v, \omega) > 0\}} - 1 \right| dv \\ (35) \quad & \leq C \int_{-1}^1 \mathbf{1}_{\{Y_n(A, s-\beta_n(\omega)v, \omega) = 0\}} dv. \end{aligned}$$

The change of variable $u = s - \beta_n(\omega)v$ yields to

$$\begin{aligned} & \int_{-1}^1 \mathbf{1}_{\{Y_n(A, s-\beta_n(\omega)v, \omega) = 0\}} dv \leq \frac{1}{\beta_n(\omega)} \int_{s-\beta_n(\omega)}^{s+\beta_n(\omega)} \mathbf{1}_{\{Y_n(A, u, \omega) = 0\}} du \\ (36) \quad & \leq \frac{1}{\beta_n(\omega)} \int_{r_2-\beta_n(\omega)}^{r_2+\beta_n(\omega)} \mathbf{1}_{\{Y_n(A, u, \omega) = 0\}} du, \end{aligned}$$

because $u \mapsto \mathbf{1}_{\{Y_n(A,u,\omega)=0\}}$ is an increasing function. Moreover, by definition of Riemann's integral of piecewise-continuous functions, we have when $n \rightarrow +\infty$,

$$\int_{r_2-\beta_n(\omega)}^{r_2+\beta_n(\omega)} \mathbf{1}_{\{Y_n(A,u,\omega)=0\}} du \sim \beta_n(\omega) \left(\mathbf{1}_{\{Y_n(A,r_2^-, \omega)=0\}} + \mathbf{1}_{\{Y_n(A,r_2^+, \omega)=0\}} \right).$$

Thus, when $n \rightarrow +\infty$,

$$(37) \quad \frac{1}{\beta_n(\omega)} \int_{r_2-\beta_n(\omega)}^{r_2+\beta_n(\omega)} \mathbf{1}_{\{Y_n(A,u,\omega)=0\}} du \sim \mathbf{1}_{\{Y_n(A,r_2^-, \omega)=0\}} + \mathbf{1}_{\{Y_n(A,r_2^+, \omega)=0\}}.$$

Futhermore, by definition of Ω_1 (see (33)),

$$(38) \quad \mathbf{1}_{\{Y_n(A,r_2,\omega)=0\}} \rightarrow 0 \quad \text{when } n \rightarrow +\infty.$$

Finally, (35), (36), (37) and (38) show that

$$\sup_{r_1 \leq s \leq r_2} \left| \int_{-1}^1 K(v) \left(\mathbf{1}_{\{Y_n(A,s-\beta_n(\omega)v,\omega)>0\}} - 1 \right) l(A, s - \beta_n(\omega)v) dv \right|$$

tends to 0 for almost all ω , since $\mathbf{P}_x(\Omega_1) = 1$.

(ii) By virtue of Lebesgue's dominated convergence theorem, we have for any $\omega \in \Omega_1$,

$$\sup_{r_1 \leq s \leq r_2} \left| \int_{-1}^1 K(u) \left(l(A, s - \beta_n(\omega)u) - l(A, s) \right) du \right| \rightarrow 0 \quad \text{when } n \rightarrow +\infty,$$

since $l(A, \cdot)$ is a continuous function.

Finally, by (34), when n goes to infinity,

$$\sup_{r_1 \leq s \leq r_2} |l_{n,\beta_n,t}^*(A, s) - l(A, s)| \rightarrow 0 \quad \mathbf{P}_x\text{-a.s.}$$

Consequently, the convergence also holds in probability. Together with (31) and (32), we obtain the expected result. \square

Proposition 4.23 states the existence of a (random) sequence $(\beta_n)_{n \geq 0}$ such that the estimator $\widehat{l}_{n,\beta_n,t}(A, \cdot)$ tends to $l(A, \cdot)$, within every compact subset $[r_1, r_2]$, with $0 < r_1 < r_2 < t$ and $t < t^*(A)$. Nevertheless, we do not obtain an explicit construction of this sequence. Furthermore, let us notice that the sequence $(\beta_n)_{n \geq 0}$ depends on t .

Finally, for any $A \in \mathcal{B}_\nu^+$, we provide a consistent estimator $\widehat{l}_{n,\beta_n,t}(A, \cdot)$ of $l(A, \cdot)$ within the interval $[0, t]$ for any $0 < t < t^*(A)$. However, we are interested in the estimation of the jump rate λ . Since λ is Lipschitz, if A is small enough, then $l(A, \cdot)$ and $\lambda(\xi, \cdot)$ are close for $\xi \in A$ (see Lemma 4.18). This leads to our main result of convergence.

Theorem 4.24. *Let \mathcal{K} be a compact subset of E and $\xi \in E$. For any $\varepsilon, \eta > 0$, there exist an integer N and a finite partition $P = (A_k)$ of \mathcal{K} such that, for any $0 < t < \min_k t^*(A_k)$, there exists for each k , a sequence $(\beta_n(A_k))_{n \geq 0}$ (depending on t) which almost surely tends to 0, such that for any $n \geq N$, for any $0 < r_1 < r_2 < t$,*

$$\mathbf{P}_\xi \left(\sup_{x \in \mathcal{K}} \sup_{r_1 \leq s \leq r_2} \left| \sum_{k=1}^{|P|} \widehat{l}_{n,\beta_n(A_k),t}(A_k, s) \mathbf{1}_{\{\widehat{\nu}_n(A_k) > \frac{1}{\sqrt{n}}\}} \mathbf{1}_{\{x \in A_k\}} - \lambda(x, s) \mathbf{1}_{\{x \in \mathcal{K}'\}} \right| > \eta \right) < \varepsilon,$$

where \mathcal{K}' is defined by

$$\mathcal{K}' = \bigcup_{\nu(A_k) > 0} A_k.$$

Proof. This is a consequence of Proposition 4.23 and Lemma 4.18. The proof is similar to the one of Theorem 4.21. \square

Remark 4.25. *If the compact subset \mathcal{K} is close to E , then, for any partition (A_k) de \mathcal{K} , the lower bound $t^*(A_k)$ is small. In this case, one estimates the jump rate on a great part of the state space, but within a small time interval. Conversely, if one chooses a small compact subset \mathcal{K} , centered in E , one estimates the jump rate on a small part of the state space, but within a larger time interval.*

5. ILLUSTRATION

Here we present a short simulation study, in order to illustrate the convergence results stated in Theorems 4.21 and 4.24. In this section, we consider a non-homogeneous marked renewal process $(X_t)_{t \geq 0}$ defined on the disk $\mathcal{D} = \{x \in \mathbf{R}^2 : \|x\|_2 \leq 1\}$ and starting from the origin $(0, 0)$. The three characteristics Q , λ and t^* are given, for any $x = (x_1, x_2) \in \mathcal{D}$, by

- for any $A \in \mathcal{B}(\mathbf{R}^2)$, $Q(x, A) = \frac{1}{K_x} \int_A \mathbf{1}_{\mathcal{D}}(y) \exp\left(-\frac{1}{8}\|y - x\|_2^2\right) dy$,
- for any $t \geq 0$, $\lambda(x, t) = \frac{|x_1| + t}{1 + \|x\|_2}$,
- $t^*(x) = 2 + \|x\|_2$,

where K_x is the normalizing constant. This transition kernel obviously satisfies Doeblin's condition. Thus, Assumption 4.1 holds for the embedded chain $(Z_n)_{n \geq 0}$. Moreover, the cumulative rate Λ associated to λ is given by

$$\Lambda(x, t) = \frac{|x_1|t}{1 + \|x\|_2} + \frac{t^2}{2 + 2\|x\|_2}.$$

We simulate a long trajectory of the process: the observation of 100000 jumps is available for estimating the jump rate λ .

In the following, we focus on the estimation of the cumulative rate $\Lambda(x, t)$ and the jump rate $\lambda(x, t)$, for $x = (0, 0)$ and $x = (0.2, 0.5)$. In both cases, we choose to approximate $\lambda(x, t)$ by the function $l(A_x, t)$, with $A_x = \{y \in \mathbf{R}^2 : \|y - x\|_2 \leq \varepsilon\}$ and $\varepsilon = 0.2$. Therefore, we only may estimate the two functions of interest within an interval smaller than $[0, 2]$. In both cases, the number of visits in $A_{(0,0)}$ and in $A_{(0.2,0.5)}$ is close to 4000 (4176 visits in $A_{(0,0)}$ and 4042 visits in $A_{(0.2,0.5)}$). In addition, the chosen bandwidth $\beta_n(A_x)$ can be written in the following way,

$$\beta_n(A_x) = \frac{1}{h_n(A_x)^\alpha},$$

where $h_n(A_x)$ denotes the (random) number of visits in A_x , and $\alpha = 1/5$. Figures 1 and 2 are given to illustrate the good behavior of our estimators of λ and Λ .

6. CONCLUDING REMARKS

In the present paper, we have developed an efficient nonparametric method for estimating both the cumulative rate and the jump rate for a general class of non-homogeneous

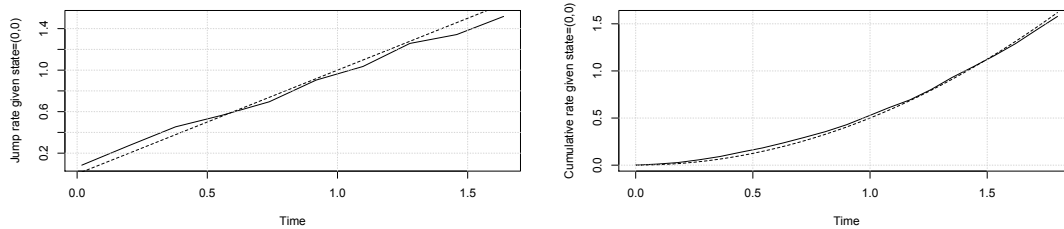


FIGURE 1. Estimation of the jump rate $\lambda(x, t)$ (left) and the cumulative rate $\Lambda(x, t)$ (right) with $x = (0, 0)$ and $0 \leq t \leq 1.8$. Estimates are drawn in solid lines, exact rates are in dashed lines.

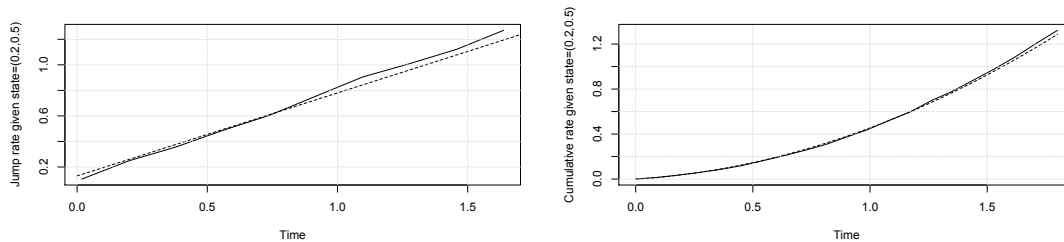


FIGURE 2. Estimation of the jump rate $\lambda(x, t)$ (left) and the cumulative rate $\Lambda(x, t)$ (right) with $x = (0.2, 0.5)$ and $0 \leq t \leq 1.8$. Estimates are drawn in solid lines, exact rates are in dashed lines.

marked renewal processes. Two inherent difficulties are related to the presence of a deterministic censorship, and the absence of conditions on the state space of the process. Furthermore, the proposed estimation procedure needs only one observation of the process within a long time interval. In addition, the assumptions which we impose are directly connected to the primitive characteristics of the process.

In this context, we proposed nonparametric estimators of the cumulative rate and the jump rate, and we proved results of uniform convergence in probability on every compact subset. As it is illustrated in Section 5, the numerical behavior of both estimators is satisfactory on finite sample size. Furthermore, the method is easy to implement and not so time-consuming. Finally, although this paper has an intrinsic interest, it is also a keystone for estimating the conditional density of the interarrival times for piecewise-deterministic Markov processes (see [6]).

REFERENCES

- [1] AALEN, O. O. *Statistical inference for a family of counting processes*. ProQuest LLC, Ann Arbor, MI, 1975. Thesis (Ph.D.)—University of California, Berkeley.
- [2] AALEN, O. O. Weak convergence of stochastic integrals related to counting processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 38, 4 (1977), 261–277.
- [3] AALEN, O. O. Nonparametric inference for a family of counting processes. *Ann. Statist.* 6, 4 (1978), 701–726.

- [4] AALEN, O. O., ANDERSEN, P. K., BORGAN, Ø., GILL, R. D., AND KEIDING, N. History of applications of martingales in survival analysis. *J. Électron. Hist. Probab. Stat.* 5, 1 (2009), 28.
- [5] ANDERSEN, P. K., BORGAN, Ø., GILL, R. D., AND KEIDING, N. *Statistical models based on counting processes*. Springer Series in Statistics. Springer-Verlag, New York, 1993.
- [6] AZAÏS, R., DUFOUR, F., AND GÉGOUT-PETIT, A. Nonparametric estimation of the jump rate for piecewise-deterministic Markov processes. *Preprint*.
- [7] BENAÏM, M., AND EL KAROUI, N. *Promenade aléatoire, Chaînes de Markov et simulations ; martingales et stratégie*. Les éd. de l'École polytechnique, Palaiseau, Paris, 2004.
- [8] BERAN, J. Nonparametric regression with randomly censored survival data, 1981. Technical report, Dept. Statist. Univ. California, Berkeley.
- [9] CHUNG, K. L. *A course in probability theory*, second ed. Academic Press, New York-London, 1974. Probability and Mathematical Statistics, Vol. 21.
- [10] COMTE, F., GAÏFFAS, S., AND GUILLOUX, A. Adaptive estimation of the conditional intensity of marker-dependent counting processes. *Ann. Inst. H. Poincaré Probab. Statist.* 47, 4 (2011), 1171–1196.
- [11] COX, D. R. Regression models and life-tables. *J. Roy. Statist. Soc. Ser. B* 34 (1972), 187–220.
- [12] DABROWSKA, D. M. Nonparametric regression with censored survival time data. *Scand. J. Statist.* 14, 3 (1987), 181–197.
- [13] DAVIS, M. H. A. *Markov models and optimization*, vol. 49 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London, 1993.
- [14] HERNÁNDEZ-LERMA, O., AND LASSERRE, J. B. Further criteria for positive Harris recurrence of Markov chains. *Proc. Amer. Math. Soc.* 129, 5 (2001), 1521–1524.
- [15] HERNÁNDEZ-LERMA, O., AND LASSERRE, J. B. *Markov chains and invariant probabilities*, vol. 211 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, 2003.
- [16] JACOBSEN, M. *Statistical analysis of counting processes*, vol. 12 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 1982.
- [17] KALLENBERG, O. *Foundations of modern probability*, second ed. Probability and its Applications (New York). Springer-Verlag, New York, 2002.
- [18] KLENKE, A. *Probability theory*. Universitext. Springer-Verlag London Ltd., London, 2008. A comprehensive course.
- [19] LI, G., AND DOSS, H. An approach to nonparametric regression for life history data using local linear fitting. *Ann. Statist.* 23, 3 (1995), 787–823.
- [20] MARTINUSSEN, T., AND SCHEIKE, T. H. *Dynamic regression models for survival data*. Statistics for Biology and Health. Springer, New York, 2006.
- [21] MCKEAGUE, I. W., AND UTIKAL, K. J. Inference for a nonlinear counting process regression model. *Ann. Statist.* 18, 3 (1990), 1172–1187.
- [22] MEYN, S., AND TWEEDIE, R. L. *Markov chains and stochastic stability*, second ed. Cambridge University Press, Cambridge, 2009.
- [23] RAMLAU-HANSEN, H. Smoothing counting process intensities by means of kernel functions. *Ann. Statist.* 11, 2 (1983), 453–466.
- [24] STUTE, W. Conditional empirical processes. *Ann. Statist.* 14, 2 (1986), 638–647.
- [25] UTIKAL, K. J. Nonparametric inference for a doubly stochastic Poisson process. *Stochastic Process. Appl.* 45, 2 (1993), 331–349.
- [26] UTIKAL, K. J. Nonparametric inference for Markovian interval processes. *Stochastic Process. Appl.* 67, 1 (1997), 1–23.

E-mail address: romain.azais@inria.fr

E-mail address: dufour@math.u-bordeaux1.fr

E-mail address: anne.petit@u-bordeaux2.fr

INRIA BORDEAUX SUD-OUEST, TEAM CQFD, FRANCE AND UNIVERSITÉ BORDEAUX, IMB, CNRS UMR 5251, 351, COURS DE LA LIBÉRATION, 33405 TALENCE CEDEX, FRANCE.